

US005552281A

United States Patent [19]  
Stashenko et al.

[11] Patent Number: 5,552,281  
[45] Date of Patent: Sep. 3, 1996

[54] HUMAN OSTEOCLAST-SPECIFIC AND -RELATED GENES

[75] Inventors: Philip Stashenko, Norfolk; Yi-Ping Li, Boston; Anne L. Wucherpfennig, Brookline, all of Mass.

[73] Assignee: Forsyth Dental Infirmary for Children, Boston, Mass.

[21] Appl. No.: 392,678

[22] Filed: Feb. 23, 1995

Related U.S. Application Data

[63] Continuation of Ser. No. 45,270, Apr. 6, 1993, abandoned.

[51] Int. Cl. 6 ..... C07H 21/04; C12N 5/10; C12N 15/70; C12Q 1/68

[52] U.S. Cl. .... 435/6; 435/69.1; 435/172.3; 435/252.3; 435/320.1; 536/23.1

[58] Field of Search ..... 435/6, 320.1, 252.3, 435/69.1, 172.3; 536/23.1

[56] References Cited

PUBLICATIONS

Blair, Harry C., et al., "Extracellular-matrix degradation at acid pH. Avian osteoclast acid collagenase isolation and characterization", *Biochemical Journal* 290(3):873-884 (15 Mar. 1993).

Tezuka, Ken-Ichi, et al., "Identification of osteoponin in isolated rabbit osteoclasts", *Biochemical and Biophysical Research Communications* 186(2):914-916 (31 Jul. 1992). Tezuka, Ken-Ichi, et al., "Molecular cloning of a possible cysteine proteinase predominantly expressed in osteoclasts", *Journal of Biological Chemistry* 269(2):1106-1108, (14 Jan. 1994).

Horton, Michael A. et al., "Monoclonal Antibodies to Osteoclastomas (Giant Cell Bone Tumors): Definition of Osteoclast-specific Cellular Antigens," *Cancer Research* 45, 5663-5669 (Nov. 1985).

Davies, John et al., "The Osteoclast Functional Antigen, Implicated in the Regulation of Bone Resorption, Is Biochemically Related to the Vitronectin Receptor," *The Journal of Cell Biology* 109, 1817-1826 (Oct. 1989).

Hayman, Alison, R. et al., "Purification and characterization of a tartrate-resistant acid phosphatase from human osteoclastomas," *Biochem. J.* 261, 601-609 (1989).

Sandberg, M. et al., "Localization of the Expression of Types I, III, and IV Collagen, TGF- $\beta$ 1 and c-fos Genes in Developing Human Calvarial Bones," *Developmental Biology* 130, 324-334 (1988).

Sandberg, M. et al., "Enhanced expression of TGF- $\beta$  and c-fos mRNAs in the growth plates of developing human long bones," *Development* 102, 461-470 (1988).

Ek-Rylander, Barbro et al., "Cloning, Sequence, and Developmental Expression of a Type 5, Tartrate-resistant, Acid Phosphatase of Rat Bone," *The Journal of Biological Chemistry* 266(36), 24684-24689 (Dec. 25, 1991).

GenBank/EMBL Sequence Search Printout, pp. 1-19 (Jun. 24, 1993).

Primary Examiner—W. Gary Jones

Assistant Examiner—Paul B. Tran

Attorney, Agent, or Firm—Hamilton, Brook, Smith & Reynolds, P.C.

ABSTRACT

The present invention relates to purified DNA sequences encoding all or a portion of an osteoclast-specific or -related gene products and a method for identifying such sequences. The invention also relates to antibodies directed against an osteoclast-specific or -related gene product. Also claimed are DNA constructs capable of replicating DNA encoding all or a portion of an osteoclast-specific or -related gene product, and DNA constructs capable of directing expression in a host cell of an osteoclast-specific or -related gene product.

5 Claims, 1 Drawing Sheet

1. AGAGAGCTT GCGGGC TGTGTTT GCGGTT GGGGGG TGGGGG  
2. GGGGGG TGGGGG TGTGTT GCGGTT GGGGGG TGGGGG  
3. CTTCAGGGG AGGGGG AGGGGG TGGGGG TGTGTT GGGGGG  
4. CTCGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
5. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
6. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
7. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
8. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
9. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
10. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
11. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
12. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
13. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
14. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
15. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
16. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
17. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
18. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
19. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
20. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
21. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
22. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
23. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
24. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
25. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
26. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
27. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
28. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
29. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
30. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
31. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
32. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
33. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
34. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
35. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
36. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
37. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
38. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
39. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
40. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
41. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
42. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
43. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
44. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
45. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
46. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
47. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
48. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
49. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
50. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
51. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
52. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
53. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
54. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
55. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
56. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
57. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
58. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
59. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
60. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
61. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
62. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
63. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
64. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
65. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
66. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
67. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
68. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
69. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
70. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
71. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
72. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
73. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
74. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
75. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
76. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
77. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
78. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
79. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
80. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
81. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
82. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
83. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
84. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
85. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
86. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
87. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
88. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
89. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
90. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
91. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
92. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
93. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
94. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
95. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
96. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
97. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
98. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
99. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
100. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
101. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
102. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
103. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
104. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
105. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
106. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
107. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
108. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
109. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
110. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
111. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
112. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
113. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
114. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
115. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
116. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
117. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
118. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
119. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
120. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
121. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
122. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
123. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
124. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
125. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
126. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
127. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
128. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
129. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
130. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
131. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
132. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
133. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
134. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
135. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
136. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
137. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
138. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
139. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
140. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
141. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
142. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
143. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
144. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
145. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
146. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
147. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
148. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
149. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
150. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
151. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
152. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
153. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
154. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
155. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
156. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
157. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
158. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
159. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
160. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
161. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
162. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
163. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
164. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
165. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
166. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
167. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
168. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
169. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
170. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
171. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
172. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
173. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
174. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
175. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
176. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
177. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
178. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
179. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
180. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
181. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
182. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
183. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
184. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
185. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
186. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
187. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
188. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
189. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
190. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
191. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
192. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
193. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
194. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
195. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
196. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
197. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
198. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
199. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
200. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
201. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
202. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
203. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
204. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
205. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
206. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
207. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
208. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
209. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
210. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
211. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
212. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
213. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
214. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
215. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
216. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
217. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
218. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
219. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
220. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
221. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
222. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
223. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
224. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
225. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
226. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
227. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
228. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
229. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
230. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
231. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
232. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
233. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
234. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
235. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
236. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
237. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
238. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
239. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
240. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
241. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
242. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
243. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
244. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
245. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
246. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
247. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
248. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
249. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
250. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
251. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
252. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
253. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
254. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
255. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
256. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
257. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
258. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
259. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
260. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
261. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
262. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
263. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
264. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
265. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
266. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
267. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
268. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
269. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
270. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
271. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
272. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
273. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
274. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
275. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
276. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
277. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
278. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
279. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
280. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
281. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
282. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
283. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
284. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
285. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
286. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
287. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
288. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
289. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
290. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
291. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
292. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
293. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
294. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
295. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
296. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
297. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
298. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
299. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
300. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
301. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
302. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
303. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
304. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
305. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
306. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
307. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
308. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
309. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
310. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
311. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
312. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
313. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
314. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
315. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG  
316. GGGGGG CGGGGG GGGGGG TGGGGG TGTGTT GGGGGG

1 AGACACCTCT GCCCTCACCA TGAGCCTCTG GCAGCCCCTG GTCCCTGGTGC TCCTGGTGCT  
61 GGGCTGCTGC TTTGCTGCC CGAGACAGCG CCAGTCCACC CTTGTGCTCT TCCCTGGAGA  
121 CCTGAGAAC AATCTCACCG ACAGGCAGCT GGCAAGAGAA TACCTGTACC GCTATGGTTA  
181 CACTCGGGTG GCAGAGATGC GTGGAGAGTC GAAATCTCTG GGGCCTGCGC TGCTGCTCT  
241 CCAGAAGCAA CTGTCCCTGC CGGAGACCGG TGAGCTGGAT AGGCCACGC TGAAGGCCAT  
301 GCGAACCCCA CGGTGCGGGG TCCCAGACCT GGGCAGATTG CAAACCTTTG AGGGCGACCT  
361 CAAGTGGCAC CACCACAACA TCACCTATTG GATCCAAAAC TACTCGGAAG ACTTGCCGCG  
421 GGCGGTGAAT GACGACGCCT TTGCCCCGCGC CTTCGCACTG TGGAGCGCGG TGACGCCGCT  
481 CACCTTCACT CGCGTGTACA GCCGGGACGC AGACATCGTC ATCCAGTTG GTGTCGCGGA  
541 GCACGGAGAC GGGTATCCCT TCGACGGGAA GGACGGGCTC CTGGCACACG CCTTCCCTCC  
601 TGGCCCCGGC ATTCAAGGGAG ACCCCCATTG CGACGATGAC GAGTGTGGT CCCTGGGCAA  
661 GGGCGTCGTG GTTCCAACTC GGTGGAAA CGCAGATGGC GCGGCGTGCC ACTTCCCCTT  
721 CATCTTCGAG GGCGCTCCT ACTCTGCTG CACCACCGAC GGTGCTCCG ACGGGTTGCC  
781 CTGGTGCAGT ACCACGGCCA ACTACGACAC CGACGACCGG TTGGCTTCT GCCCGAGCGA  
841 GAGACTCTAC ACCCGGGACG GCAATGCTGA TGGGAAACCC TGCCAGTTTC CATTCACTT  
901 CCAAGGCCAA TCCTACTCCG CCTGCACAC GGACGGTCGC TCCGACGGCT ACCGCTGGTG  
961 CGCCACCACC GCCAACTACG ACCGGGACAA GCTCTCGGC TTCTGCCGA CCCGAGCTGA  
1021 CTCGACGGTG ATGGGGGCA ACTCGGCGGG GGAGCTGTGC GTCTCCCT TCACTTCCCT  
1081 GGGTAAGGAG TACTCGACCT GTACCGACGA CGGCGCGGA GATGGCGCC TCTGGTGC  
1141 TACCACTCG AACCTTGACA GCGACAAGAA GTGGGGCTTC TGCCCGGACC AAGGATAACAG  
1201 TTGTTCTC GTGGCGCGC ATGAGITCGG CCACGCGCTG GGCTTAGATC ATTCTCACT  
1261 GCCGGAGGCG CTCATGTACC CTATGTACCG CTTCACTGAG GGGCCCCCT TGCATAAGGA  
1321 CGACGTGAAT GGCACTCCGGC ACCCTCTATGG TCCCTGCCCT GAACCTGAGC CACGGCCTCC  
1381 AACCACCAAC ACACCGCAGC CCACGGCTCC CCCGACGGTC TGCCCGACCG GACCCCCCAC  
1441 TGTCCACCCC TCAGAGCGCC CCACAGCTGG CCCCACAGGT CCCCCCTCAG CTGGCCCCAC  
1501 AGGTCCCCCCC ACTGCTGGCC CTTCTACGGC CACTACTGTG CCTTGAGTC CGGTGGACGA  
1561 TGCCTGCAAC GTGAACATCT TCGACGCCAT CGCGGAGATT GGGAACCGC TGTATTTGTT  
1621 CAAGGATGGG AAGTACTGGC GATTCTCTGA GGGCAGGGGG AGCCGGCCGC AGGGCCCCCTT  
1681 CCTTATCGCC GACAAGTGGC CGCGCGCTGCC CGCGAAGCTG GACTCGGTCT TTGAGGAGCC  
1741 GCTCTCCAAG AAGCTTTCT TCTTCTCTGG GCGCCAGGTG TGGGTGTACA CAGGCGCGTC  
1801 GGTGCTGGGC CGGAGGCAGTC TGGACAAGCT GGGCCTGGGA GCCGACGTGG CCCAGGTGAC  
1861 CGGGGCCCTC CGGAGTGGCA GGGGAAGAT GCTGCTGTTG AGCAGGGCGC GCCTCTGGAG  
1921 GTTCGACGTG AAGGCGCAGA TGGTGGATCC CGGGAGCGCC AGCGAGGTGG ACCGGATGTT  
1981 CCCCCGGGTG CCTTGGACA CGCACGACGT CTTCCAGTAC CGAGAGAAAG CCTATTTCTG  
2041 CCAGGACCGC TTCTACTGGC CGGTGAGTTG CGGGAGTGAG TTGAACCAAGG TGGACCAAGT  
2101 GGGCTACGTG ACCTATGACA TCCTGCACTG CCCTGAGGAC TAGGGCTCCC GTCCTGCTT  
2161 GCAGTGCCAT GTAAATCCCC ACTGGGACCA ACCCTGGGA AGGAGCCAGT TTGCGGGATA  
2221 CAAACTGGTA TTCTGTTCTG GAGGAAGGG AGGAGTGGAG GTGGGCTGGG CCCTCTCTTC  
2281 TCACCTTGT TTTTGTGG AGTGTCTAATAAAACTTGG ATTCTCTAAC CTTT

1

**HUMAN OSTEOCLAST-SPECIFIC AND  
-RELATED GENES**

**RELATED APPLICATION**

This application is a continuation of application Ser. No. 08/045,270 filed on Apr. 6, 1993 now abandoned.

**BACKGROUND OF THE INVENTION**

Excessive bone resorption by osteoclasts contributes to the pathology of many human diseases including arthritis, osteoporosis, periodontitis, and hypercalcemia of malignancy. During resorption, osteoclasts remove both the mineral and organic components of bone (Blair, H. C., et al., *J. Cell Biol.* 102:1164 (1986)). The mineral phase is solubilized by acidification of the sub-osteoclastic lacuna, thus allowing dissolution of hydroxyapatite (Vaes, G., *Clin. Orthop. Relat.* 231:239 (1988)). However, the mechanism(s) by which type I collagen, the major structural protein of bone, is degraded remains controversial. In addition, the regulation of osteoclastic activity is only partly understood. The lack of information concerning osteoclast function is due in part to the fact that these cells are extremely difficult to isolate as pure populations in large numbers. Furthermore, there are no osteoclastic cell lines available. An approach to studying osteoclast function that permits the identification of heretofore unknown osteoclast-specific or -related genes and gene products would allow identification of genes and gene products that are involved in the resorption of bone and in the regulation of osteoclastic activity. Therefore, identification of osteoclast-specific or -related genes or gene products would prove useful in developing therapeutic strategies for the treatment of disorders involving aberrant bone resorption.

**SUMMARY OF THE INVENTION**

The present invention relates to isolated DNA sequences encoding all or a portion of osteoclast-specific or -related gene products. The present invention further relates to DNA constructs capable of replicating DNA encoding osteoclast-specific or -related gene products. In another embodiment, the invention relates to a DNA construct capable of directing expression of all or a portion of the osteoclast-specific or -related gene product in a host cell.

Also encompassed by the present invention are prokaryotic or eukaryotic cells transformed or transfected with a DNA construct encoding all or a portion of an osteoclast-specific or -related gene product. According to a particular embodiment, these cells are capable of replicating the DNA construct comprising the DNA encoding the osteoclast-specific or -related gene product, and, optionally, are capable of expressing the osteoclast-specific or -related gene product. Also claimed are antibodies raised against osteoclast-specific or -related gene products, or portions of these gene products.

The present invention further embraces a method of identifying osteoclast-specific or -related DNA sequences and DNA sequences identified in this manner. In one embodiment, cDNA encoding osteoclast is identified as follows: First, human giant cell tumor of the bone was used to 1) construct a cDNA library; 2) produce <sup>32</sup>P-labelled cDNA to use as a stromal cell\*, osteoclast\* probe, and 3) produce (by culturing) a stromal cell population lacking osteoclasts. The presence of osteoclasts in the giant cell tumor was confirmed by histological staining for the osteoclast marker, type 5 acid phosphatase (TRAP). The cultured stromal cell population did not contain osteoclasts. The cultured stromal cell population was then used to produce a stromal cell\*, osteoclast\* <sup>32</sup>P-labelled cDNA probe.

2

clast marker, type 5 tartrate-resistant acid phosphatase (TRAP) and with the use of monoclonal antibody reagents.

The stromal cell population lacking osteoclasts was produced by dissociating cells of a giant cell tumor, then growing and passaging the cells in tissue culture until the cell population was homogeneous and appeared fibroblastic. The cultured stromal cell population did not contain osteoclasts. The cultured stromal cells were then used to produce a stromal cell\*, osteoclast\* <sup>32</sup>P-labelled cDNA probe.

10 The cDNA library produced from the giant cell tumor of the bone was then screened in duplicate for hybridization to the cDNA probes: one screen was performed with the giant cell tumor cDNA probe (stromal cell\*, osteoclast\*), while a duplicate screen was performed using the cultured stromal cell cDNA probe (stromal cell\*, osteoclast\*). Hybridization to a stromal\*, osteoclast\* probe, accompanied by failure to hybridize to a stromal\*, osteoclast\* probe indicated that a clone contained nucleic acid sequences specifically expressed by osteoclasts.

20 In another embodiment, genomic DNA encoding osteoclast-specific or -related gene products is identified through known hybridization techniques or amplification techniques. In one embodiment, the present invention relates to a method of identifying DNA encoding an osteoclast-specific or -related protein, or gene product, by screening a cDNA library or a genomic DNA library with a DNA probe comprising one or more sequences selected from the group consisting of the DNA sequences set out in Table I (SEQ ID NOs: 1-32). Finally, the present invention relates to an osteoclast-specific or related protein encoded by a nucleotide sequence comprising a DNA sequence selected from the group consisting of the sequences set out in Table I, or their complementary strands.

35

**BRIEF DESCRIPTION OF FIG. 1**

The FIG. 1 shows cDNA sequence (SEQ ID NO: 33) of human gelatinase B, and highlights those portions of the sequence represented by the osteoclast-specific or -related cDNA clones of the present invention.

40

**DETAILED DESCRIPTION OF THE  
INVENTION**

45 As described herein, Applicant has identified osteoclast-specific or osteoclast-related nucleic acid sequences. These sequences were identified as follows: Human giant cell tumor of the bone was used to 1) construct a cDNA library; 2) produce <sup>32</sup>P-labelled cDNA to use as a stromal cell\*, osteoclast\* probe, and 3) produce (by culturing) a stromal cell population lacking osteoclasts. The presence of osteoclasts in the giant cell tumor was confirmed by histological staining for the osteoclast marker, type 5 acid phosphatase (TRAP). In addition, monoclonal antibody reagents were used to characterize the multinucleated cells in the giant cell tumor, which cells were found to have a phenotype distinct from macrophages and consistent with osteoclasts.

50 The stromal cell population lacking osteoclasts was produced by dissociating cells of a giant cell tumor, then growing the cells in tissue culture for at least five passages. After five passages the cultured cell population was homogeneous and appeared fibroblastic. The cultured population contained no multinucleated cells at this point, tested negative for type 5 acid phosphatase, and tested variably alkaline phosphatase positive. That is, the cultured stromal cell population did not contain osteoclasts. The cultured stromal

55

60

65

cells were then used to produce a stromal cell<sup>+</sup>, osteoclast<sup>-</sup><sup>32</sup>P-labelled cDNA probe.

The cDNA library produced from the giant cell tumor of the bone was then screened in duplicate for hybridization to the cDNA probes: one screen was performed with the giant cell tumor cDNA probe (stromal cell<sup>+</sup>, osteoclast<sup>-</sup>), while a duplicate screen was performed using the cultured stromal cell cDNA probe (stromal cell<sup>+</sup> osteoclast<sup>-</sup>). Clones that hybridized to the giant cell tumor cDNA probe (stromal<sup>+</sup>, osteoclast<sup>-</sup>), but not to the stromal cell cDNA probe (stromal<sup>+</sup>, osteoclast<sup>-</sup>), were assumed to contain nucleic acid sequences specifically expressed by osteoclasts.

As a result of the differential screen described herein, DNA specifically expressed in osteoclast cells characterized as described herein was identified. This DNA, and equivalent DNA sequences, is referred to herein as osteoclast-specific or osteoclast-related DNA. Osteoclast-specific or -related DNA of the present invention can be obtained from sources in which it occurs in nature, can be produced recombinantly or synthesized chemically; it can be cDNA, genomic DNA, recombinantly-produced DNA or chemically-produced DNA. An equivalent DNA sequence is one which hybridizes, under standard hybridization conditions, to an osteoclast-specific or -related DNA identified as described herein or to a complement thereof.

Differential screening of a human osteoclastoma cDNA library was performed to identify genes specifically expressed in osteoclasts. Of 12,000 clones screened, 195 clones were identified which are either uniquely expressed in osteoclasts, or are osteoclast-related. These clones were further identified as osteoclast-specific, as evidenced by failure to hybridize to mRNA derived from a variety of unrelated human cell types, including epithelium, fibroblasts, lymphocytes, myelomonocytic cells, osteoblasts, and neuroblastoma cells. Of these, 32 clones contain novel cDNA sequences which were not found in the GenBank database.

A large number of cDNA clones obtained by this procedure were found to represent 92 kDa type IV collagenase (gelatinase B; E.C. 3.4.24.35) as well as tartrate resistant acid phosphatase. In situ hybridization localized mRNA for gelatinase B to multinucleated giant cells in human osteoclastomas. Gelatinase B immunoreactivity was demonstrated in giant cells from 8/8 osteoclastomas, osteoclasts in normal bone, and in osteoclasts of Paget's disease by use of a polyclonal antisera raised against a synthetic gelatinase B peptide. In contrast, no immunoreactivity for 72 kDa type IV collagenase (gelatinase A; E.C. 3.4.24.24), which is the product of a separate gene, was detected in osteoclastomas or normal osteoclasts.

The present invention has utility for the production and identification of nucleic acid probes useful for identifying osteoclast-specific or -related DNA. Osteoclast-specific or -related DNA of the present invention can be used to produce osteoclast-specific or -related gene products useful in the therapeutic treatment of disorders involving aberrant bone resorption. The osteoclast-specific or -related sequences are also useful for generating peptides which can then be used to produce antibodies useful for identifying osteoclast-specific or -related gene products, or for altering the activity of osteoclast-specific or -related gene products. Such antibodies are referred to as osteoclast-specific antibodies. Osteoclast-specific antibodies are also useful for identifying osteoclasts. Finally, osteoclast-specific or -related DNA sequences of the present invention are useful in gene therapy. For example, they can be used to alter the

expression in osteoclasts of an aberrant osteoclast-specific or -related gene product or to correct aberrant expression of an osteoclast-specific or -related gene product. The sequences described herein can further be used to cause osteoclast-specific or related gene expression in cells in which such expression does not ordinarily occur, i.e., in cells which are not osteoclasts.

#### Example 1—Osteoclast cDNA Library Construction

Messenger RNA (mRNA) obtained from a human osteoclastoma ('giant cell tumor of bone'), was used to construct an osteoclastoma cDNA library. Osteoclastomas are actively bone resorptive tumors, but are usually non-metastatic. In cryostat sections, osteoclastomas consist of ~30% multinucleated cells positive for tartrate resistant acid phosphatase (TRAP), a widely utilized phenotypic marker specific in vivo for osteoclasts (Minkin, *Calcif. Tissue Int.* 34:285-290 (1982)). The remaining cells are uncharacterized 'stromal' cells, a mixture of cell types with fibroblastic/mesenchymal morphology. Although it has not yet been definitively shown, it is generally held that the osteoclasts in these tumors are non-transformed, and are activated to resorb bone in vivo by substance(s) produced by the stromal cell element.

Monoclonal antibody reagents were used to partially characterize the surface phenotype of the multinucleated cells in the giant cell tumors of long bone. In frozen sections, all multinucleated cells expressed CD68, which has previously been reported to define an antigen specific for both osteoclasts and macrophages (Horton, M. A. and M. H. Helfrich, In *Biology and Physiology of the Osteoclast*, B. R. Rifkin and C. V. Gay, editors, CRC Press, Inc. Boca Raton, Fla., 33-54 (1992)). In contrast, no staining of giant cells was observed for CD11b or CD14 surface antigens, which are present on monocytes/macrophages and granulocytes (Arnaout, M. A. et al. *J. Cell. Physiol.* 137:305 (1988); Haziol, A. et al. *J. Immunol.* 141:547 (1988)). Cytocentrifuge preparations of human peripheral blood monocytes were positive for CD68, CD11b, and CD14. These results demonstrate that the multinucleated giant cells of osteoclastomas have a phenotype which is distinct from that of macrophages, and which is consistent with that of osteoclasts.

Osteoclastoma tissue was snap frozen in liquid nitrogen and used to prepare poly A<sup>+</sup> mRNA according to standard methods. cDNA cloning into a pcDNAII vector was carried out using a commercially-available kit (Librarian, InVitrogen). Approximately 2.6×10<sup>6</sup> clones were obtained, >95% of which contained inserts of an average length 0.6 kB.

#### Example 2—Stromal Cell mRNA Preparation

A portion of each osteoclastoma was snap frozen in liquid nitrogen for mRNA preparation. The remainder of the tumor was dissociated using brief trypsinization and mechanical disaggregation, and placed into tissue culture. These cells were expanded in Dulbecco's MEM (high glucose, Sigma) supplemented with 10% newborn calf serum (MA Bioproducts), gentamycin (0.5 mg/ml), L-glutamine (2 mM) and non-essential amino acids (0.1 mM) (Gibco). The stromal cell population was passaged at least five times, after which it showed a homogenous, fibroblastic looking cell population that contained no multinucleated cells. The stromal cells were mononuclear, tested negative acid phosphatase, and tested variably alkaline phosphatase positive. These findings indicate that propagated stromal cells (i.e., stromal cells that

are passaged in culture) are non-osteoclastic and non-activated.

**Example 3—Identification of DNA Encoding Osteoclastoma-Specific or -Related Gene Products by Differential Screening of an Osteoclastoma cDNA Library**

A total of 12,000 clones drawn from the osteoclastoma cDNA library were screened by differential hybridization, using mixed <sup>32</sup>P labelled cDNA probes derived from (1) giant cell tumor mRNA (stromal cell<sup>+</sup>, OC<sup>+</sup>), and (2) mRNA from stromal cells (stromal cell<sup>+</sup>, OC<sup>-</sup>) cultivated from the same tumor. The probes were labelled with <sup>32</sup>PdCTP by random priming to an activity of ~10<sup>6</sup>CPM/μg. Of these 12,000 clones, 195 gave a positive hybridization signal with giant cell (i.e., osteoclast and stromal cell) mRNA, but not with stromal cell mRNA. Additionally, these clones failed to hybridize to cDNA produced from mRNA derived from a variety of unrelated human cell types including epithelial cells, fibroblasts, lymphocytes, myelomonocytic cells, osteoblasts, and neuroblastoma cells. The failure of these clones to hybridize to cDNA produced from mRNA derived from other cell types supports the conclusion that these clones are either uniquely expressed in osteoclasts, or are osteoclast-related.

The osteoclast (OC) cDNA library was screened for differential hybridization to OC cDNA (stromal cell<sup>+</sup>, OC<sup>+</sup>) and stromal cell cDNA (stromal cell<sup>+</sup>, OC<sup>-</sup>) as follows:

NYTRAN filters (Schleicher & Schuell) were placed on agar plates containing growth medium and ampicillin. Individual bacterial colonies from the OC library were randomly picked and transferred, in triplicate, onto filters with pre-pledged grids and then onto a master agar plate. Up to 200 colonies were inoculated onto a single 90-mm filter/plate using these techniques. The plates were inverted and incubated at 37° C. until the bacterial inoculates had grown (on the filter) to a diameter of 0.5–1.0 mm.

The colonies were then lysed, and the DNA bound to the filters by first placing the filters on top of two pieces of Whatman 3 MM paper saturated with 0.5N NaOH for 5 minutes. The filters were neutralized by placing on two pieces of Whatman 3 MM paper saturated with 1M Tris-HCl, pH 8.0 for 3–5 minutes. Neutralization was followed by incubation on another set of Whatman 3 MM papers saturated with 1M Tris-HCl, pH 8.0/1.5M NaCl for 3–5 minutes. The filters were then washed briefly in 2×SSC.

DNA was immobilized on the filters by baking the filters at 80° C. for 30 minutes. Filters were best used immediately, but they could be stored for up to one week in a vacuum jar at room temperature.

Filters were prehybridized in 5–8 ml of hybridization solution per filter, for 2–4 hours in a heat sealable bag. An additional 2 ml of solution was added for each additional filter added to the hybridization bag. The hybridization

buffer consisted of 5×SSC, 5×Denhardt's solution, 1% SDS and 100 μg/ml denatured heterologous DNA.

Prior to hybridization, labeled probe was denatured by heating in 1×SSC for 5 minutes at 100° C., then immediately chilled on ice. Denatured probe was added to the filters in hybridization solution, and the filters hybridized with continuous agitation for 12–20 hours at 65° C.

After hybridization, the filters were washed in 2×SSC/0.2% SDS at 50°–60° C. for 30 minutes, followed by washing in 0.2×SSC/0.2% SDS at 60° C. for 60 minutes.

The filters were then air dried and autoradiographed using an intensifying screen at -70° C. overnight.

**Example 4—DNA Sequencing of Selected Clones**

Clones reactive with the mixed tumor probe, but unreactive with the stromal cell probe, are expected to contain either osteoclast-related, or *in vivo* 'activated' stromal-cell-related gene products. One hundred and forty-four cDNA clones that hybridized to tumor cell cDNA, but not to stromal cell cDNA, were sequenced by the dideoxy chain termination method of Sanger et al. (Sanger F., et al. *Proc. Natl. Acad. Sci. USA* 74:S463 (1977)) using sequenase (US Biochemical). The DNASIS (Hitachi) program was used to carry out sequence analysis and a homology search in the GenBank/EMBL database.

Fourteen of the 195 tumor<sup>+</sup> stromal<sup>-</sup> clones were identified as containing inserts with a sequence identical to the osteoclast marker, type 5 tartrate-resistant acid phosphatase (TRAP) (GenBank accession number J04430 M19534). The high representation of TRAP positive clones also indicates the effectiveness of the screening procedure in enriching for clones which contain osteoclast-specific or related cDNA sequences.

Interestingly, an even larger proportion of the tumor<sup>+</sup> stromal<sup>-</sup> clones (77/195; 39.5%) were identified as human gelatinase B (macrophage-derived gelatinase) (Wilhelm, S. *M. J. Biol. Chem.* 264:17213 (1989)), again indicating high expression of this enzyme by osteoclasts. Twenty-five of the gelatinase B clones were identified by dideoxy sequence analysis; all 25 showed 100% sequence homology to the published gelatinase B sequence (Genbank accession number J05070). The portions of the gelatinase B cDNA sequence covered by these clones is shown in the FIGURE (SEQ ID NO: 33). An additional 52 gelatinase B clones were identified by reactivity with a <sup>32</sup>P-labelled probe for gelatinase B.

Thirteen of the sequenced clones yielded no readable sequence. A DNASIS search of GenBank/EMBL databases revealed that, of the remaining 91 clones, 32 clones contain novel sequences which have not yet been reported in the databases or in the literature. These partial sequences are presented in Table I. Note that three of these sequences were repeats, indicating fairly frequent representation of mRNA related to this sequence. The repeat sequences are indicated by <sup>a</sup>, <sup>b</sup> superscripts (Clones 198B, 223B and 32C of Table I).

TABLE I

PARTIAL SEQUENCES OF 32 NOVEL OC-SPECIFIC OR -RELATED EXPRESSED GENES (cDNA CLONES)

34A (SEQ ID NO: 1)					
1 GCAAATACT	AAGTTTATTC	CTTGGATTTC	TAGTGAGAGC	TGTTGAATT	GGTGATGTCA
61 AATGTTTCTA	GGGTTTTTT	AGTTTGTTC	TATTAAAAAA	TTTAATTAT	TATGCTATAG
121 GTGATACTCT	CTTGAATAA	ACCTATAATA	GAAAATAGCA	GCAGACAACA	
4B (SEQ ID NO: 2)					
1 GTCTCAACCT	GCATATCCTA	AAAATGTCAA	AATGCTGCAT	CTGTTAACG	TCGGGGTAGG

TABLE I-continued

## PARTIAL SEQUENCES OF 32 NOVEL OC-SPECIFIC OR -RELATED EXPRESSED GENES (cDNA CLONES)

61	GGG					
12B (SEQ ID NO: 3)						
1	CTTCCCTCTC	TTGCTTCCT	TTCCCCAAGCA	GAGGTGCTCA	CTCCATGCC	ACCGCCACCA
61	CAGGCCACA	GGGAGTACTG	CCAGACTACT	GCTGATGTC	TCTTAAGGCC	CAGGGAGTCT
121	CAACAGCTG	GTGGTGAATG	CTGCTGGCA	CGGGACCCCC	CCC	
28B (SEQ ID NO: 4)						
1	TTTATTTGT	AAATATATGT	ATTACATCCC	TAGAAAAAAGA	ATGCCAGAT	TTTCCCTCCT
61	GTGTGTTTC	GTCTTGCTTC	TTCATGGTCC	ATGATGCCAG	CTGAGGTTGT	CAGTACAATG
121	AAACCAACT	GGCGGGAATG	AAGCAATTAA	TTCTGCGATT	TTCTGCGTC	TTT
37B (SEQ ID NO: 5)						
1	GGCTGGACAT	GGGTGCCCCTC	CACGTGCTTC	ATATCCCCAG	GCACACTCTG	GCCTCAGGTT
61	TTGCCCTGGC	CATGTCATCT	ACCTGGAGTG	GGGCTCCCCC	TTCCTCACCC	TTGAATCAA
121	AGCCACTTGT	TTAGGCAGG	AITTCAGAAGA	CCACTCATCA	CAATAAAAAA	TATTTGAAA
181	ACAAAAAAA	AAAAAAA				
55B (SEQ ID NO: 6)						
1	TIGACAAAGC	TGTTTATTC	CACCAATAAA	TAGTATATGG	TGATTGGGT	TTCTATTTAT
61	AAGAGTAGTG	GCTATTATAT	GGGGTATCAT	GTGATGTC	ATAAAATAGTT	CATATCTACT
121	TAATTGCT	TC				
60B (SEQ ID NO: 7)						
1	GAAGAGAGTT	GTATGTACAA	CCCCAACAGG	CAAGGCAGCT	AAATGCAGAG	GGTACAGAGA
61	GATCCCGAGG	GAATT				
86B (SEQ ID NO: 8)						
1	GGATGAAAC	ATGTAGAAGT	CCAGAGAAAA	ACAATTTAA	AAAAAGGTGG	AAAAGTTACG
61	GCAAAACCTGA	GATTTCAGCA	TAAAATCTTT	AGTTAGAAGT	GAAGAGAAAGA	AGAGGGAGGC
121	TGGTTGGTGT	TGCACTGATC	AATAGGTAT	C		
87B (SEQ ID NO: 9)						
1	TTCTTGTACT	TTAGAACACT	ATGAATAGGG	AAAAAAAGAAA	AAACTGTICA	AAATAAAAATG
61	TAGGAGCCGT	GCTTTGGAA	TGCTTGAGTG	AGGAGCTCAA	CAAGTCCTCT	CCCAAGAAAG
181	CAATGATAAA	ACTTGACAAA	A			
98B (SEQ ID NO: 10)						
1	ACCCATTCT	AAACATTTT	ACTGTAAAAT	TTTGGTCAA	AGTTCTAACG	TTAACACAT
61	CTCAAAGAAT	AGAGGCAATA	TATAGCCAT	CTTACTAGAC	ATACAGTATT	AAACTGGACT
121	GAATATGAGG	ACAAGCTCTA	GTGGTCATTA	AAACCCCTCAG	AA	
110B (SEQ ID NO: 11)						
1	ACATATATA	ACAGCAITCA	TTTGGCCAAA	ATCTACACGT	TTGAGAAC	CTACTGTATA
61	TAAGTGGGA	ATGTATCAAG	TATAGACTAT	GAAGTGCAA	ATAACAAGTC	AAGGTTAGAT
121	TAACCTTTT	TTTTTACATT	ATAAAATTAA	CTTGTTT		
118B (SEQ ID NO: 12)						
1	CCAAATTCT	CTGGAAATCCA	TCCTCCCTCC	CATCACCATA	GCCTCGAGAC	GTCAATTCTG
61	TTTGACTACT	CCAGC				
133B (SEQ ID NO: 13)						
1	AACTAACCTC	CTCGGACCCC	TGCTCACTC	ATTTACACCA	ACCACCAAC	TATCTATAAA
61	CCTGAGCCAT	GGCCATCCCT	TATGACCGC	GCAGTGATTA	TAGGCTTCG	CTCTAAAGATA
121	AAAT					
140B (SEQ ID NO: 14)						
1	TTTATTATTC	TTTTTTATG	TTAGCTTAGC	CATGAAAAT	TTACTGGTGA	AGCAGTTAAAT
61	AAAACACACA	TCCCATTGAA	GGGTTTTGTA	CATTTCAGTC	CTTACAATA	ACAAAGCAAT
121	GATAAACCCG	GCACGCTCTG	ATAGGAATT	C		
144B (SEQ ID NO: 15)						
1	CGTGACACAA	ACATGCATTC	GTTTTATCA	TAACACAGCC	TGGTTCTTA	AAACAATACA
61	AACAGCATGT	TCATCAGCAG	GAAGCTGCC	CTGGGAGGG	GGGCC	
198B* (SEQ ID NO: 16)						
1	ATAGGTAGA	TTCTCATTC	CGGGACTAGT	TAGCTTAAAG	CACCTAGAG	GAATAGGTA
61	ATCTGACTTC	TCACCTTC	AGTTCCCTCT	TATATCTCA	AGGTAGAAAT	GTCTATGTTT
121	TCTACTCCAA	TTCAATAATC	TATTCTATAAG	TCTTGGTAC	AAGTTACATG	ATAAAAAGAA
181	ATGTGATTG	TCTTCCCTTC	TTTGACTT	TRAATAAAAG	TATTTATCTC	CTGTCTACAG
241	TTTAAT					
212B (SEQ ID NO: 17)						
1	GTCCAGTATA	AAGGAAACCG	TIAACTCGGT	AGCTAGAGG	ATGTAATA	TCTTTATGT
61	CCTCTAGATA	AAACACCCGA	TTAACAGATG	TTAACCTTT	AIGTTTGT	TTGCTTTAAA
121	AAAGGCTTC	TACACATTAG	CTCCAGCTAA	AAAGACACAT	TGAGAGCTTA	GAGGATAGTC
181	TCTGGAGC					
223B* (SEQ ID NO: 18)						
1	GCACTTGGAA	GGGAGTTGGT	GTGCTATT	TGAACAGAT	GTGGTGTAC	TGAAGATTGTC
61	TGTTCAATTG	CCCCATTGTT	TTGTGCTTC	AATGATCTT	CCTACTTTC	TTCCTCTCCAC
121	CCATGAACCT	TTTCACTGTG	GCCATCAAGG	ACTTCTCTGA	CAGCTTGTGT	ACTCTTAGGC
181	TAAGAGATGT	GACTACAGCC	TGCCCCGTAC	TG		
241B (SEQ ID NO: 19)						
1	TGTTAGTTT	TAGGAAGGCC	TGTCTTCTGG	GAATGAGGTT	TATTAGTCCA	CTTCTGGAG
61	CTAGACGTCC	TATAGTTAGT	CACTGGGGAT	GGTGAAGAG	GGAGAAAGGG	AAGGGCGAAG
121	GGAAAGGGCTC	TTTGCTAGTA	TCTCCATTTC	TAGAAGATGG	TTTATGATGT	AACCACAGGT
181	CTATATGAGC	ATAGTAAGGC	TGT			
32C* (SEQ ID NO: 20)						
1	CCTATTCTG	ATCTGACTT	TGGACAAGGC	CCTTCAGCCA	GAAGACTGAC	AAAGTCATCC
121	TCCGTCCTACC	AGAGCGTGC	CTTGTGATCC	TAATAAAGC	TTCATCTCG	GCTGTGCCTT
181	GGGTGGAAGG	GGCAGGATTC	TGCAGCTGT	TTTCATTA	TTCCTAA	TTTCATT

TABLE I-continued

## PARTIAL SEQUENCES OF 32 NOVEL OC-SPECIFIC OR -RELATED EXPRESSED GENES (cDNA CLONES)

34C (SEQ ID NO: 21)					
1 CGGAGGTTAG	GTGTGTTTAT	TCTGTACAA	ATCATTACAA	AACCAAGTCT	GGGGCAGTCA
61 CCCCCCCCAC	CCATCACCCC	AGTGCATGG	CTAGCTGCTG	GCCTTT	
47C (SEQ ID NO: 22)					
1 TTAGTTCACT	CAAAGCAGGC	AACCCCTTT	GGCACTGCTG	CCACTGGGT	CATGGCGTT
61 GTGCCAGCTG	GGGAGGTTTC	CCCAACACCC	TCCTCTGCTT	CCCTGTGTGT	CGGGGTCTCA
121 GGAGCTGACC	CAGACTGGA				
65C (SEQ ID NO: 23)					
1 GCTGAATGTT	TAAGAGAGAT	TTTGGTCTTA	AAGGCTTCAT	CATGAAAGTG	TACATGCATA
61 TGCAAGTGTG	AATTACGTGG	TATGGATGGT	TGTTTGTGTTA	TTAACTAAAG	ATGTACAGCA
121 AACTGCCCT	TTAGAGTCCT	CTTAATATTG	ATGTCCTAAC	ACTGGGTCTG	CTTATGC
79C (SEQ ID NO: 24)					
1 GCGCACTGGGA	TATGGAATCC	AGAAGGGAAA	CAAGCACTGG	ATAATTTAAA	ACAGCTGGGG
61 AGAAAACCTGG	GAAAACAAAG	GATAJATOCT	CATGGCTCGA	AATAAGAACAA	ACGCTGTGG
121 CATTGCAAC	CTGGCCAGCT	TCOCCAAGAT	GTGACTCCAG	CCAGAAA	
84C (SEQ ID NO: 25)					
1 GCCAGGGGGG	ACCGCTTTA	TTCCCTCTCT	GCCTCAGAGG	TCAGGAAGGA	GGTCTGGCAG
61 GACCTGCACT	GGGCCCTAGT	CATCTGTGGC	AGCGAAGGTG	AAGGGACTCA	CCTTGTGCGC
121 CGTGCCTGAG	TAGAACCTGT	TCTGGAATTC	C		
86C (SEQ ID NO: 26)					
1 AACTCTTC	CACTCTGGTA	TITTTAGTTT	AACAATATAT	GTGTTGTC	TTGAAAATA
61 GTTCATATCA	ATTCTATTC	AGCTGTCTCA	TTCTTTTTTT	AATGGTCATA	TACAGTAGTA
121 TTCAATTATA	AGAAATATTC	CTAATACTTT	TTAAAAA		
87C (SEQ ID NO: 27)					
1 GGATAAGAAA	GAAGGCCCTGA	GGCCTAGGGG	CCRGGGCTGG	CCTGCGTCTC	AGTOCTGGGA
61 CGCAGCAGCC	CGCACAGGT	GAGAGGGGCA	CTTCCCTCTTG	CTTAGGTTGG	TGAGGATCTG
121 GTCTCTGGTG	GCCGTGGAG	AGCCACAAA			
88C (SEQ ID NO: 28)					
1 CTGACCTTCG	AGACCTTGAC	CTGAGGCCCG	ATACCTACTG	CCGCTATGAC	TCGGTCAGCG
61 TGTICAACGG	AGCGCTGAGC	GACCGACTCG	GTGGGGAAAGT	TCTGCGCGA	T
89C (SEQ ID NO: 29)					
1 ATCCCTGGCT	GTGGATAGTG	CTTTTGTGTA	CCAAATGCTC	OCTCCTTAAG	GTATAGGGC
61 TCCCTGAGTT	TGGAGATGTG	GAAGTACTAC	TTAACGTCT	GTCTGCTTG	GCTGTOGTTA
121 TCGTTTCTG	GTGATGTGTT	GCTAACATA	AGAATAC		
101C (SEQ ID NO: 30)					
1 GGCTGGGCAT	CCCTCTCTC	CTCCATCCCC	ATACATCACC	AGGTCTAAATG	TTTACAAACG
61 GTGCCAGCCC	GGCTCTGAAG	CCAAGGGCCG	TCCGTGCCAC	GGTGGCTGTG	AGTATTCTC
121 CGTTAGCTT	CCCATAAGGT	TGGAGTACT	GC		
112C (SEQ ID NO: 31)					
1 CCAACTCTTA	CCCGGATACA	GACCCACAGA	GTGCCATCCC	TGAGAGACCA	GACOGCTCCC
161 CAATACTCTC	CTAAAATAAA	CATGAAGCAC			
114C (SEQ ID NO: 32)					
1 CATGGATGAA	TGTCTCATGG	TGGGAAGGAA	CATGGTACAT	TTC	

<sup>a</sup>Repeated 3 times<sup>b</sup>Repeated 2 times

Sequence analysis of the OC<sup>+</sup> stromal cell<sup>-</sup> cloned DNA sequences revealed, in addition to the novel sequences, a number of previously-described genes. The known genes identified (including type 5 acid phosphatase, gelatinase B, cystatin C (13 clones), Alu repeat sequences (11 clones), creatine kinase (6 clones) and others) are summarized in Table II. In situ hybridization (described below) directly demonstrated that gelatinase B mRNA is expressed in multinucleated osteoclasts and not in stromal cells. Although gelatinase B is a well-characterized protease, its expression at high levels in osteoclasts has not been previously described. The expression in osteoclasts of cystatin C, a cysteine protease inhibitor, is also unexpected. This finding has not yet been confirmed by in situ hybridization. Taken together, these results demonstrate that most of these identified genes are osteoclast-expressed, thereby confirming the effectiveness of the differential screening strategy for identifying DNA encoding osteoclast-specific or -related gene products. Therefore, novel genes identified by this method have a high probability of being OC-specific or related.

In addition, a minority of the genes identified by this screen are probably not expressed by OCs (Table II). For example, type III collagen (6 clones), collagen type I (1 clone), dermatansulfate (1 clone), and type VI collagen (1

45

clone) are more likely to originate from the stromal cells or from osteoblastic cells which are present in the tumor. These cDNA sequences survive the differential screening process either because the cells which produce them in the tumor in vivo die out during the stromal cell propagation phase, or because they stop producing their product in vitro. These clones do not constitute more than 5-10% of the all sequences selected by differential hybridization.

TABLE II  
SEQUENCE ANALYSIS OF CLONES ENCODING KNOWN SEQUENCES FROM AN OSTEOCLASTOMA cDNA LIBRARY

Clones with Sequence Homology to Collagenase Type IV	23 total
Clones with Sequence Homology to Type 5 Taurate Resistant Acid Phosphatase	14 total
Clones with Sequence Homology to Cystatin C	13 total
Clones with Sequence Homology to Alu-repeat Sequences	11 total
Clones with Sequence Homology to Creatine Kinase	6 total
Clones with Sequence Homology to	6 total

TABLE II-continued

SEQUENCE ANALYSIS OF CLONES ENCODING KNOWN SEQUENCES FROM AN OSTEOCLASTOMA cDNA LIBRARY	
Type III Collagen	5 total
Clones with Sequence Homology to MHC Class I $\gamma$ Invariant Chain	3 total
Clones with Sequence Homology to MHC Class II $\beta$ Chain	10 total
One or Two Clone(s) with Sequence Homology to Each of the Following:	
$\alpha 1$ collagen type I	
$\gamma$ interferon inducible protein	
osteopontin	
Human chondroitin/dermatan sulfate	15
$\alpha$ globin	
$\beta$ glucuronidase/sphingolipid activator	
Human CAPL protein (Ca binding)	
Human EST 01024	
Type VI collagen	
Human EST 00553	

12  
UTP digoxigenin labelled cRNA probes.

TABLE III

S	In Situ HYBRIDIZATION USING PROBES DERIVED FROM NOVEL SEQUENCES		
	Clone	Osteoclasts	Stromal Cells
4B	+	+	
28B*	+	-	
37B	+	+	
86B	-	-	
87B	-	-	
88C	+	+	
98B	+	+	
118B*	+	-	
140B*	+	-	
198B*	+	-	
212B*	+	-	
Gelatinase B*	+	-	

\*OC-expressed, as indicated by reactivity with antisense probe and lack of reactivity with sense probe on OCs only.

## Example 5—In situ Hybridization of OC-Expressed Genes

In situ hybridization was performed using probes derived from novel cloned sequences in order to determine whether the novel putative OC-specific or -related genes are differentially expressed in osteoclasts (and not expressed in the stromal cells) of human giant cell tumors. Initially, in situ hybridization was performed using antisense (positive) and sense (negative control) cRNA probes against human type IV collagenase/gelatinase B labelled with  $^{35}$ S-UTP.

A thin section of human giant cell tumor reacted with the antisense probe resulted in intense labelling of all OCs, as indicated by the deposition of silver grains over these cells, but failed to label the stromal cell elements. In contrast, only minimal background labelling was observed with the sense (negative control) probe. This result confirmed that gelatinase B is expressed in human OCs.

In situ hybridization was then carried out using cRNA probes derived from 11/32 novel genes, labelled with digoxigenin UTP according to known methods.

The results of this analysis are summarized in Table III. Clones 28B, 118B, 140B, 198B, and 212B all gave positive reactions with OCs in frozen sections of a giant cell tumor, as did the positive control gelatinase B. These novel clones therefore are expressed in OCs and fulfill all criteria for OC-relatedness. 198B is repeated three times, indicating relatively high expression. Clones 4B, 37B, 88C and 98B produced positive reactions with the tumor tissue; however the signal was not well-localized to OCs. These clones are therefore not likely to be useful and are eliminated from further consideration. Clones 86B and 87B failed to give a positive reaction with any cell type, possibly indicating very low level expression. This group of clones could still be useful but may be difficult to study further. The results of this analysis show that 5/11 novel genes are expressed in OCs, indicating that ~50% of novel sequences likely to be OC-related.

To generate probes for the in situ hybridizations, cDNA derived from novel cloned osteoclast-specific or -related cDNA was subcloned into a BlueScript II SK(-) vector. The orientation of cloned inserts was determined by restriction analysis of subclones. The T7 and T3 promoters in the BlueScriptII vector was used to generate  $^{35}$ S-labelled ( $^{35}$ S-UTP 850 Ci/mmol, Amersham, Arlington Heights, Ill.), or

25  
30  
35  
40  
45  
50  
55  
60  
65

In situ hybridization was carried out on 7 micron cryostat sections of a human osteoclastoma as described previously (Chang, L.-C. et al. *Cancer Res.* 49:6700 (1989)). Briefly, tissue was fixed in 4% paraformaldehyde and embedded in OCT (Miles Inc., Kankakee, Ill.). The sections were rehydrated, postfixed in 4% paraformaldehyde, washed, and pretreated with 10 mM DTT, 10 mM iodoacetamide, 10 mM N-ethylmaleimide and 0.1 triethanolamine-HCl. Prehybridization was done with 50% deionized formamide, 10 mM Tris-HCl, pH 7.0, 1x Denhardt's, 500 mg/ml tRNA, 80 mg/ml salmon sperm DNA, 0.3M NaCl, mM EDTA, and 100 mM DTT at 45° C. for 2 hours. Fresh hybridization solution containing 10% dextran sulfate and 1.5 ng/ml  $^{35}$ S-labelled or digoxigenin labelled RNA probe was applied after heat denaturation. Sections were coverslipped and then incubated in a moistened chamber at 45°–50° C. overnight. Hybridized sections were washed four times with 50% formamide, 2x SSC, containing 10 mM DTT and 0.5% Triton X-100 at 45° C. Sections were treated with RNase A and RNase T1 to digest single-stranded RNA, washed four times in 2x SSC/10 mM DTT.

In order to detect  $^{35}$ S-labelling by autoradiography, slides were dehydrated, dried, and coated with Kodak NTB-2 emulsion. The duplicate slides were split, and each set was placed in a black box with desiccant, sealed, and incubated at 4° C. for 2 days. The slides were developed (4 minutes) and fixed (5 minutes) using Kodak developer D19 and Kodak fixer. Hematoxylin and eosin were used as counterstains.

In order to detect digoxigenin-labelled probes, a Nucleic Acid Detection Kit (Boehringer-Mannheim, Cal. #1175041) was used. Slides were washed in Buffer 1 consisting of 100 mM Tris/150 mM NaCl, pH 7.5, for 1 minute. 100  $\mu$ l Buffer 2 was added (made by adding 2 mg/ml blocking reagent as provided by the manufacturer) in Buffer 1 to each slide. The slides were placed on a shaker and gently swirled at 20° C.

Antibody solutions were diluted 1:100 with Buffer 2 (as provided by the manufacturer). 100  $\mu$ l of diluted antibody solution was applied to the slides and the slides were then incubated in a chamber for 1 hour at room temperature. The slides were monitored to avoid drying. After incubation with antibody solution, slides were washed in Buffer 1 for 10 minutes, then washed in Buffer 3 containing 2 mM levamisole for 2 minutes.

After washing, 100  $\mu$ l color solution was added to the slides. Color solution consisted of nitroblue/tetrazolium salt

13

(NBT) (1:225 dilution) 4.5  $\mu$ l, 5-bromo-4-chloro-3-indolyl phosphate (1:285 dilution) 3.5  $\mu$ l, levamisole 0.2 mg in Buffer 3 (as provided by the manufacturer) in a total volume of 1 ml. Color solution was prepared immediately before use.

After adding the color solution, the slides were placed in a dark, humidified chamber at 20° C. for 2-5 hours and monitored for color development. The color reaction was stopped by rinsing slides in TE Buffer.

The slides were stained for 60 seconds in 0.25% methyl green, washed with tap water, then mounted with water-based Permount (Fisher).

#### Example 6—Immunohistochemistry

15 Immunohistochemical staining was performed on frozen and paraffin embedded tissues as well as on cytocentrifuge preparations (see Table IV). The following antibodies were used: polyclonal rabbit anti-human gelatinase antibodies; Ab110 for gelatinase B; monoclonal mouse anti-human CD68 antibody (clone KP1) (DAKO, Denmark); Mol (anti-CD11b) and Mo2 (anti-CD14) derived from ATCC cell lines HB CRL 8026 and TIB 228/HB44. The anti-human gelatinase B antibody Ab110 was raised against a synthetic peptide with the amino acid sequence EALMYPMYRFTEGPPPLHK (SEQ ID NO: 34), which is specific for human gelatinase B (Corcoran, M. L. et al. *J. Biol. Chem.* 267:515 (1992)).

20 Detection of the immunohistochemical staining was achieved by using a goat anti-rabbit glucose oxidase kit (Vector Laboratories, Burlingame Calif.) according to the manufacturer's directions. Briefly, the sections were rehydrated and pretreated with either acetone or 0.1% trypsin. Normal goat serum was used to block nonspecific binding. Incubation with the primary antibody for 2 hours or overnight (Ab110:1/500 dilution) was followed by either a glucose oxidase labeled secondary anti-rabbit serum, or, in the case of the mouse monoclonal antibodies, were reacted with purified rabbit anti-mouse Ig before incubation with the secondary antibody.

25 Paraffin embedded and frozen sections from osteoclastomas (GCT) were reacted with a rabbit antiserum against gelatinase B (antibody 110) (Corcoran, M. L. et al. *J. Biol. Chem.* 267:515 (1992)), followed by color development with glucose oxidase linked reagents. The osteoclasts of a giant cell tumor were uniformly strongly positive for gelatinase B, whereas the stromal cells were unreactive. Control sections reacted with rabbit preimmune serum were negative. Identical findings were obtained for all 8 long bone giant cell tumors tested (Table IV). The osteoclasts present in three out of four central giant cell granulomas (CCG) of the mandible were also positive for gelatinase B expression. These neoplasms are similar but not identical to the long bone giant cell tumors, apart from their location in the jaws (Shafer, W. G. et al. *Textbook of Oral Pathology*, W. B. Saunders Company, Philadelphia, pp. 144-149 (1983)). In contrast, 30 the multinucleated cells from a peripheral giant cell tumor, which is a generally non-resorptive tumor of oral soft tissue,

10

15

20

25

30

35

40

45

50

14

were unreactive with antibody (Shafer, W. G. et al., *Textbook of Oral Pathology*, W. B. Saunders Company, Philadelphia, pp. 144-149 (1983)).

Antibody 110 was also utilized to assess the presence of gelatinase B in normal bone (n=3) and in Paget's disease, in which there is elevated bone remodeling and increased osteoclastic activity. Strong staining for gelatinase B was observed in osteoclasts both in normal bone (mandible of a 2 year old), and in Paget's disease. Staining was again absent in controls incubated with preimmune serum. Osteoblasts did not stain in any of the tissue sections, indicating that gelatinase B expression is limited to osteoclasts in bone. Finally, peripheral blood monocytes were also reactive with antibody 110 (Table IV).

TABLE IV

## DISTRIBUTION OF GELATINASE B IN VARIOUS TISSUES

Sample	Antibodies tested
	Ab 110 gelatinase B
GCT frozen (n = 2)	
giant cells	+
stromal cells	-
GCT paraffin (n = 6)	
giant cells	+
stromal cells	-
central CCG (n = 4)	
giant cells	+ (34)
stromal cells	-
peripheral GCT (n = 4)	
giant cells	-
stromal cells	-
Paget's disease (n = 1)	
osteoclasts	+
osteoblasts	-
normal bone (n = 3)	
osteoclasts	+
osteoblasts	-
monocytes (cytocentrifuge)	+

Distribution of gelatinase B in multinucleated giant cells, osteoclasts, osteoblasts and stromal cells in various tissues. In general, paraffin embedded tissues were used for these experiments; exceptions are indicated.

## Equivalents

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments described herein. Such equivalents are intended to be encompassed by the following claims.

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## (111) NUMBER OF SEQUENCES: 34

5,552,281

15

16

-continued

---

( 2 ) INFORMATION FOR SEQ ID NO:1:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 170 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double
- ( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:1:

GCAAATATCT AAGTTTATTG CTTGGATTTC TAGTGACAGC TGTGAATTG GGTGATGTCA	60
AATGTTCTA GGGTTTTTT AGTTTGTGTT TATTGAAAAAA TTTAATTATT TATGCTATAG	120
GTGATATTCT CTTGAATAA ACCTATAATA GAAAATAGCA GCAGACAAACA	170

( 2 ) INFORMATION FOR SEQ ID NO:2:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 63 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double
- ( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GTGTCACCT GCATATCCTA AAAATGTCAA AATGCTGCAT CTGGTTAATG TCGGGGTAGG	60
GGG	63

( 2 ) INFORMATION FOR SEQ ID NO:3:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 163 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double
- ( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:3:

CTTCCCTCTC TTGCTTCCCT TTCCCCAGCA GAGGTGCTCA CTCCATGCC ACCGCCACCA	60
CAGGCCCCACA GGGAGTACTG CCAGACTACT GCTGATGTTC TCTTAAGGCC CAAGGAGTCT	120
CAACCAACTG GTGGTGAATG CTGCCCTGGCA CGGGACCCCC CCC	163

( 2 ) INFORMATION FOR SEQ ID NO:4:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 173 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double
- ( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:4:

TTTTATTTGT AAATATATGT ATTACATCCC TAGAAAAAGA ATCCCAGGAT TTTCCCTCCT	60
GTGTGTTTC GTCTTCTTC TTCATGGTCC ATGATGCCAG CTGAGGTTGT CAGTACAATG	120
AAACCAAACG GCCGGCATGG AAGCAGATTA TTCTOCCATT TTTCCAGGTC TTT	173

( 2 ) INFORMATION FOR SEQ ID NO:5:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 197 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double

5,552,281

17

18

-continued

#### ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( z i ) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GGCTGGAACAT	GGGTGCCCTC	CACGTECCCTC	ATATCCCCAG	GCACACTCTG	GCCTCAGGTT	60
TTGCCCTGGC	CATGTCACTCT	ACCTGGAGTG	GGCCCTCCCC	TTCTTCAGCC	TTGAATCAAA	120
AGCCACTTTG	TTAGGCGAGG	ATTTCCCAGA	CCACTCATCA	CATTAaaaaaa	TATTTTGAaa	180
ACaaaaaaaaaa	Aaaaaaaaaa					197

( 2 ) INFORMATION FOR SEQ ID NO:6

### ( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 132 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double

TGACAAAGC TGTTTATTC CACCAATAA TAGTATATGG TGATTGGGT TTCTATTTAT 60  
AAGAGTAGTG GCTATTATAT GGGGTATCAT GTTGATGCTC ATAAATAGTT CATATCTACT 120  
TAATTTGCCT TC 132

( 2 ) INFORMATION FOR SEQ ID NO:7:

### ( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 75 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double
- ( D ) TOPOLOGY: linear

( i : i ) MOLECULE TYPE: DNA (economic)

( z : ) SEQUENCE DESCRIPTION: SEQ ID NO:7:

**GAAGAGAGTT** GATGTCACAA CCCCAACAGG CAAGGCAGCT AAATOCAGAG GGTACAGAGA 60  
**GATCCCGAGG** GAATT 75

( 2 ) INFORMATION FOR SEQ ID NO:8:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 151 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double
- ( D ) TOPOLOGY: linear

( i : i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```

GGATGAAAC ATGTAGAAGT CCAGAGAAAA ACAATTAA AAAAGGTGG AAAAGTTACG 60
GCAAACTGA GATTTCAGCA TAAAATCTT AGTTAGAAGT GAGAGAAAAGA AGAGGGAGGC 120
TGGTTGCTGT TGCACTATC AATAGGTTAT C 151

```

( 2 ) INFORMATION FOR SEQ ID NO:9:

### ( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 141 base pair
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

5,552,281

19

20

-continued

---

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TTCTTGTATCT TTAGAACACT ATGAAATAGGG AAAAAAAGAAA AAACCTGTTCA AAATAAAAATG 60  
TAGGAGCCGT GCTTTGGAA TGCTTGAGTG AGGAGCTAA CAAGTCCTCT CCCAAGAAAG 120  
CAATGATAAA ACTTGACAAA A 141

( 2 ) INFORMATION FOR SEQ ID NO:10:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 162 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:10:

ACCCATTCTC AACAAATTTT ACTGTAAAAT TTTTGGTCAA AGTTCTAAGC TTAATCACAT 60  
CTCAAAGAAAT AGAGGCAATA TATAAGCCCAT CTTACTAGAC ATACAGTATT AAACCTGGACT 120  
GAATATGAGG ACAAGCTCTA GTGGTCATTA AACCCCTCAG AA 162

( 2 ) INFORMATION FOR SEQ ID NO:11:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 157 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:11:

ACATATATTA ACAGCATTCA TTTGGCCAAA ATCTACACGT TTGTAGAACAT CTACTGTATA 60  
TAAAGTGGGA ATGTATCAAG TATAGACTAT GAAAGTGCAG ATAACAAGTC AAGGTTAAGAT 120  
TAACITTTTT TTTTACATT ATAAAATTAA CTTGTTT 157

( 2 ) INFORMATION FOR SEQ ID NO:12:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 75 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CCAAATTCTC CTGGAATCCA TCCTCCCTCC CATCACCCATA GCCTCGAGAC GTCATTTCTG 60  
TTTGACTACT CCAGC 75

( 2 ) INFORMATION FOR SEQ ID NO:13:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 124 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AACCTAACCTC CTCGGACCCC TGCCTCACTC ATTTACACCA ACCACCCAAAC TATCTATAAA 60  
CCTGAGCCAT GGCCATCCCT TATGAGCGGC GCAGTGATTA TAGGCTTTCG CTCTAAGATA 120

5,552,281

21

22

-continued

AAAT

124

( 2 ) INFORMATION FOR SEQ ID NO:14:

- ( i ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 151 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: double  
( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ATTATTATTC TTTTTTATG TTAGCTTAGC CATGCAAAAT TTACTGGTGA AGCAAGTTAAT	60
AAAAACACACA TCCCATTGAA GGGTTTTGTA CATTTCAGTC CTTACAAATA ACAAAAGCAAT	120
GATAAAACCCG GCACGTCCTG ATAGGAAATT C	151

( 2 ) INFORMATION FOR SEQ ID NO:15:

- ( i ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 105 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: double  
( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CGTGACACAA ACATGCATTC GTTTTATTCA TAAAACAGCC TGGTTTCCTA AAACAAATACA	60
AACAGCATGT TCATCAOCAG GAAGCTGGCC GTGGGCAGGG GGGCC	105

( 2 ) INFORMATION FOR SEQ ID NO:16:

- ( i ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 246 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: double  
( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:16:

ATAAGTTAGA TTCTCATTCA CGGGACTAGT TAGCTTTAAG CACCCTAGAG GACTAGGGTA	60
ATCTGACTTC TCACTTCCTA AGTTCCCTCT TATATCCTCA AGGTAGAAAT GTCTATGTTT	120
TCTACTCCAA TTCATAAAATC TATTCTATAAG TCTTTGGTAC AAOTTACATG ATAAAAAGAA	180
ATGTGATTTG TCTTCCCTTC TTTGCACCTT TGAAATAAAG TATTATCTC CTOTCTACAG	240
TTTAAT	246

( 2 ) INFORMATION FOR SEQ ID NO:17:

- ( i ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 188 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: double  
( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:17:

GTCCCACTATA AAGGAAAGCG TTAAGTCGTT AAGCTAGAGG ATTGTAAATA TCTTTTATGTT	60
CCTCTAGATA AAACACCCGA TTAACAGATG TTAACCTTTT ATGTTTTGAT TTGCTTTAAA	120
AATGGCCCTTC TACACATTAG CTCCAGCTAA AAAGACACAT TGAGAGCTTA GAGGATAGTC	180

5,552,281

23

24

-continued

TCTGGAGC

188

( 2 ) INFORMATION FOR SEQ ID NO:18:

- ( i ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 212 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: double  
( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:18:

GCACCTGGAA	GGGAGTTGGT	GTGCTATTTC	TGAACAGAT	GTGGTGATAAC	TGAGATTGTC	60
TGTTCA	GTTTCCCC	CAGTTTGT	TTGTGCTTCA	AATGATCCTT	CCTACTTTGC	120
CCATGACCTT	TTTCACTGTG	GCCATCAAGG	ACTTTCTGA	CAGCTTGTGT	ACTCTTAGGC	180
TAAGAGATGT	GA	CTACAGCC	TGCCCCCTGAC	TG		212

( 2 ) INFORMATION FOR SEQ ID NO:19:

- ( i ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 203 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: double  
( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TGTTA	TTTT	TAGGAAGGCC	TOTCTTCTGG	GA	GTGAGGTT	TATTAGTCCA	CTTCTTOGAG	60
CTAGAC	GTCC	TATAGTTAGT	CACTGGGGAT	GGTGA	AAAGAG	GGAGAAGAGG	AAGGGCGAAG	120
GGAAGGGCTC	TTTGCTA	GT	TCTCCATTTC	TAGAAGATGG	TTT	AGATGAT	AACCACAGGT	180
CTATATGAGC	ATAGTAAGGC	TGT						203

( 2 ) INFORMATION FOR SEQ ID NO:20:

- ( i ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 177 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: double  
( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CCTATTTCTG	ATCCTGACTT	TGGACAAGGC	CCTTCAGCCA	GAAGACTGAC	AAAGTCATCC	60
TCCGTCTACC	AOAGCGTGCA	CTTGTGATCC	TAAAATAAGC	TTCATCTCCG	GCTGTGCCTT	120
GGGTGGAAGG	GGCAGGATTC	TGCACCTGCT	TTTGCAATTTC	TCTTCCTAAA	TTTCATT	177

( 2 ) INFORMATION FOR SEQ ID NO:21:

- ( i ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 106 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: double  
( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CGGAGCGTAG	GTGTGTTTAT	TCCTGTACAA	ATCATTACAA	AACCAAGTCT	GGGGCAGTCA	60
CGGCCCCCAC	CCATCACCCC	AOTGCAATGG	CTAGCTGCTG	GCCTTT		106

5,552,281

25

26

-continued

---

( 2 ) INFORMATION FOR SEQ ID NO:22:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 139 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double
- ( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TTAGTTCACT CAAAGCAGGC AACCCCTTT GOCACTGCTG CCACGTGGGT CATGGCGGTT	60
GTGGCAGCTG GGGAGGTTTC CCCAACACCC TCCCTCTGCTT CCCTGTGTGT CGGGGTCTCA	120
GGAGCTGACC CAGAGTGG	139

( 2 ) INFORMATION FOR SEQ ID NO:23:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 177 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double
- ( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GCTGAATGTT TAAGAGAGAT TTTGGTCTTA AAGGCTTCAT CATGAAAGTG TACATGCCATA	60
TCCAAGTGTG ATTACGTGG TATGGATGGT TGCTTGTGTTA TTAACTAAAG ATGTACAGCA	120
AACTGCCCCGT TTAGAGTCCT CTTAATATTG ATGTCTAAC ACTGGGTCTG CTTATG	177

( 2 ) INFORMATION FOR SEQ ID NO:24:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 167 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double
- ( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GCCAGTGGGA TATGAAATCC AGAAGGGAAA CAAGCACTGG ATAATTAAAA ACAGCTGGGG	60
AGAAAAACTGG OGAAACAAAG GATATATCCT CATGGCTCGA AATAAGAACCA ACGCCCTGTGG	120
CATTGCCAAC CTGGCCAGCT TCCCCAAGAT GTGACTCCAG CCAGAAA	167

( 2 ) INFORMATION FOR SEQ ID NO:25:

( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 151 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double
- ( D ) TOPOLOGY: linear

( ii ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GCCAGGGCGG ACCGTCCTTA TTCCCTCTCCT GCCTCAGAGG TCAGGAAGGA GGTCTGGCAG	60
GACCTGCACT GGGCCCTAGT CATCTGTGGC AGCGAAGGTG AAGGGACTCA CCTTGTCGCC	120
CGTGCCTGAG TAGAACTTGT TCTGGAATTC C	151

( 2 ) INFORMATION FOR SEQ ID NO:26:

( i ) SEQUENCE CHARACTERISTICS:

5,552,281

27

28

-continued

- 
- ( A ) LENGTH: 156 base pairs
  - ( B ) TYPE: nucleic acid
  - ( C ) STRANDEDNESS: double
  - ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:26:

AAC TCT TTCA CACT CTG TGA TTT TTA GGT TTT AACA AAT ATAT GTG TTG TGTC TTGG AAATTA 60  
GTT CAT ATAT CA ATT CAT ATTG AGC TGT CTCA TTCT TTT TTT AAT GGT CATA TAC AGT TAG TA 120  
TTCA ATTATA AGA AAT ATATC CTA AAT ACT TT TTAAAAA 156

( 2 ) INFORMATION FOR SEQ ID NO:27:

- ( i ) SEQUENCE CHARACTERISTICS:
- ( A ) LENGTH: 150 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GGATAAGAAA GAAGGGCCTGA GGGCTAGGGGG CGGGGGCCTGG CCTGCCTCTC AGTCCCTGGGA 60  
CGCAGCAGCC CGCACAGGTT GAGAGGGGGCA CTTCCCTCTTG CTTAGGTTGG TGAGGATCTG 120  
GTCCTGTTG GCCGGTGGAG AGCCACAAAAA 150

( 2 ) INFORMATION FOR SEQ ID NO:28:

- ( i ) SEQUENCE CHARACTERISTICS:
- ( A ) LENGTH: 212 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GCAC TTT GGAA GGGAGTTGGT GTGCTATTTT TGAAGCAGAT GTGGT GATAAC TGAGATTGTC 60  
TGTT CA GTTTT CCCCATTTGT TTGTGCTTCA AATGATCCTT CCTACTTTGC TTCTCTCCAC 120  
CCATGACCTT TTTC ACTGTG GCCATCAAGG ACTTT CCTGA CAGCTTG TGT ACTCTTAGGC 180  
TAAGAGATGT GACTACAGCC TCCCCCTGAC TG 212

( 2 ) INFORMATION FOR SEQ ID NO:29:

- ( i ) SEQUENCE CHARACTERISTICS:
- ( A ) LENGTH: 157 base pairs
- ( B ) TYPE: nucleic acid
- ( C ) STRANDEDNESS: double
- ( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:29:

ATCCCCTGGCT GTGGATAGTG CTTTTGTGTA GCAAATGCTC CCTCCTTAAG GTTATAGGGC 60  
TCCCCTGAGTT TGGGAGTGTG GAAAGTACTAC TTAACTGTCT GTCCCTGCTTG GCTGTCGTTA 120  
TCGTTTTCTG GTGATGTTGT OCTAACAAATA AGAATAC 157

( 2 ) INFORMATION FOR SEQ ID NO:30:

- ( i ) SEQUENCE CHARACTERISTICS:
- ( A ) LENGTH: 152 base pairs
- ( B ) TYPE: nucleic acid

5,552,281

29

30

-continued

( C ) STRANDEDNESS: double  
( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:30:

GGCTGGGCAT CCCTCTCCTC CTCCATCCCC ATACATCACC AGGTCTAATG TTTACAAACG	60
GTOCCAGCCC GGCTCTGAAG CCAAGGGCCC TCCOTCCAC QGTGCGTGT AGTATTCTC	120
CCTTAGCTTT CCCATAAGGT TGGAGTATCT GC	152

( 2 ) INFORMATION FOR SEQ ID NO:31:

( i ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 90 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: double  
( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:31:

CCAACTCCTA CCOCGATACA GACCCACAGA GTGCCATCCC TGAGAGACCA GACCGCTCCC	60
CAATACTCTC CTAAAATAAA CATGAAGCAC	90

( 2 ) INFORMATION FOR SEQ ID NO:32:

( i ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 43 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: double  
( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CATGGATGAA TGTCTCATGG TGGGAAGGAA CATGGTACAT TTC	43
---	----

( 2 ) INFORMATION FOR SEQ ID NO:33:

( i ) SEQUENCE CHARACTERISTICS:  
( A ) LENGTH: 2333 base pairs  
( B ) TYPE: nucleic acid  
( C ) STRANDEDNESS: double  
( D ) TOPOLOGY: linear

( i i ) MOLECULE TYPE: DNA (genomic)

( x i ) SEQUENCE DESCRIPTION: SEQ ID NO:33:

AGACACCTCT GCCCTCACCA TGAGCCCTTG GCAGCCCCTG GTCCCTGGTGC TCCTGGTGC	60
GGGCTGCTGC TTTGCTGCC CCAGACAGCG CCAGTCCACCC CTTGCTGCTCT TCCCTGGAGA	120
CCTGAGAACCC AATCTCACCG ACAGGCAGCT GGCAAGAAGAA TACCTGTACCC GCTATGGTTA	180
CACTCGGGTG GCAGAGATGC GTGGAGAGTC GAAATCTCTG GGGCCTGCGC TGCTGCTTCT	240
CCAGAAGCAA CTOTCCCTGC CGAGACCGG TGAAGCTGGAT AGCGCCACGC TGAAGGCCAT	300
GCGAACCCCA CGGTGCGGGG TCCCAGACCT GGGCAGATTC CAAACCTTG AGGGCGACCT	360
CAAOTGOCAC CACCAACACA TCACCTATTG GATCCAAAAAC TACTCGGAAG ACTTGCCCGCG	420
GCGCGTGAATT GACGACGCCCT TTGCCCCGCGC CTTCGCACTG TGGAGCGCGG TGACGCCGCT	480
CACCTTCACT COCOTGTACA GCGGGGACGC AGACATCGTC ATCCAGTTG GTGTCGCGGA	540
GCACGGAGAC GGGTATCCCT TCGACGGGAA GGACGGGCTC CTGGCACACG CCTTCTCC	600
TOGCCCGGCGC ATTCAAGGGAG ACGCCCATTT CGACGATGAC GAGTTGTGGT CCCGGGCAA	660

5,552,281

31

32

-continued

GGGCOTCGTG	GTTCCAACTC	GGTTTGGAAA	CGCAGATGGC	GGGGCCTGCC	ACTTCCCCCTT	720
CATCTTCGAG	GGCCGCTCCT	ACTCTGCCCTG	CACCAACGAC	GGTCGCTCCG	ACGGGTTGCC	780
CTGGTGCAGT	ACCAACGGCCA	ACTACGACAC	CGACGGACCGG	TTTGGCTTCG	GCCCCAOGGA	840
GAGACTCTAC	ACCCGGGGACG	CGAAATGCTGA	T0000AAACCC	TOCCAGTTTC	CATTGATCTT	900
CCAAGGCCAA	TCCTACTCCG	CCTGCACCCAC	GGACGGTCGC	TCCGACGGCT	ACCGCTGGTG	960
CGCCACCAAC	GCCAACTACG	ACCGGGACAA	GCTCTCGGC	TTCTGCCCGA	CCCGAGCTGA	1020
CTCGACGCGT	ATGGGGGGCGA	ACTCGGGGGGG	GGAGCTGTGC	GTCTTCCCCCT	TCACTTTCTT	1080
GGGTAAGGAG	TACTCGACCT	GTACCAGCGA	GGGCGCGCGA	GATGGGCGCC	TCTGOTGC	1140
TACCACTCTG	AACCTTGACA	CGCACAAGAA	GTGGGCGTTC	TGCCCCGGACC	AAOGGATACAG	1200
TTTGTTCCTC	GTGGCGGC	ATGAGTTCCG	CCACGGCTG	GGCTTAGATC	ATTCCCTCAGT	1260
OCCGGAGGCG	CTCATGTACC	CTATGTACCG	CTTCACTGAG	GGGGCCCCCT	TGCATAAOGA	1320
CGACGTGAAT	GGCAGTCCGGC	ACCTCTATGG	TCTCTGCCCT	GAACCTGAGC	CACCGCCCTCC	1380
AAACCACCAAC	ACACCGCAGC	CCACGGCTCC	CCCAGCGTC	TGCCCCACCG	GACCCCCCAC	1440
TGTCCACCCC	TCAGAGCGCC	CCACAGCTGG	CCCCACAGGT	CCCCCTCAG	CTGGGCCCCAC	1500
AGGTCCCCCCC	ACTGCTGGCC	CTTCTACGGC	CACTACTGTG	CCCTTGAGTC	CGGTGGACGA	1560
TGCCTGCAAC	GTGAACATCT	TCGACGCCAT	CGCGGAGATT	GGGAACCAGC	TGTATTTGTT	1620
CAAGGATGGG	AACTACTGGC	GATTCTCTOA	GGGCAGGGGG	AGCCGGCCGC	AGGGCCCCCTT	1680
CCTTATCGCC	GACAAGTGGC	CCGCGCTGCC	CCGCAAGCTG	GACTCGGTCT	TTGAGGAGCC	1740
GCTCTCCAAG	AAGCTTTCT	TCTTCTCTGG	CGGCCAGGTG	TGGGTGTACA	CAGGCGCGTC	1800
GGTGTCTGGC	CCOAGCGTC	TGGACAAGCT	GGGCTGGGA	GCCGACGTGG	CCCAGGTGAC	1860
CGGGGCCCTC	CGGAGTGGCA	GGGGGAAGAT	GCTGCTGTC	AGCGGGCGGC	GCCTCTGGAG	1920
GTTCGACGTG	AAGGCACAGA	TGGTGGATCC	CCGGAGCGCC	AOCGAGGTGG	ACCGGATGTT	1980
CCCCGGGGTG	CCTTGGACA	CGCACGACGT	CTTCCAGTAC	CGAGAGAAAAG	CTTATTTCTG	2040
CCAGGACCGC	TTCTACTGGC	CGCTGAGTTC	CCGGAGTGAG	TTGAACCAGG	TOGACCAAOT	2100
GGGCTACGTG	ACCTATGACA	TCTGCGAGTG	CCCTGAGGAC	TAGGGCTCCC	GTCCCTGCTTT	2160
GCAGTGCAT	GTAAATCCCC	ACTGGGACCA	ACCTGGGGGA	AOGAGCCAGT	TTGCCGGATA	2220
CAAACCTGGTA	TTCTGTTCTG	GAGGAAAGGG	AGGAOTGGAG	GTGGGCGTGGG	CCCTCTCTTC	2280
TCACCTTTGT	TTTTGTTGG	AGTGTTCCTA	ATAAAACTTGG	ATTCTCTAAC	CTTT	2334

## ( 2 ) INFORMATION FOR SEQ ID NO:34:

## ( i ) SEQUENCE CHARACTERISTICS:

- ( A ) LENGTH: 18 amino acids
- ( B ) TYPE: amino acid
- ( C ) STRANDEDNESS: single
- ( D ) TOPOLOGY: unknown

## ( ii ) MOLECULE TYPE: peptide

## ( iii ) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Glu	Ala	Lys	Met	Tyr	Pro	Met	Tyr	Arg	Phe	Thr	Glu	Gly	Pro	Pro	Lys
1															15

His Lys

We claim:

1. An isolated osteoclast-specific or -related DNA sequence, or its complementary sequence, the DNA sequence comprising a nucleic acid sequence selected from the group consisting of:

a) DNA sequences set forth in the group consisting of SEQ ID NOS. 12, 14, 16 and 17, or their complementary strands; and

**33**

- b) DNA sequences which hybridize under standard conditions to the DNA sequences defined in a).
- 2. A DNA construct capable of replicating, in a host cell, osteoclast-specific or -related DNA, said construct comprising:
  - a) a DNA sequence of claim 1; and
  - b) sequences, in addition to said DNA sequence, necessary for transforming or transfecting a host cell, and for replicating, in a host cell, said DNA sequence.
- 3. A DNA construct capable of replicating and expressing, in a host cell, osteoclast-specific or -related DNA, said construct comprising:

5

**34**

- a) a DNA sequence of claim 2; and
- b) sequences, in addition to said DNA sequence, necessary for transforming or transfecting a host cell, and for replicating and expressing, in a host cell, said DNA sequence.
- 4. A cell stably transformed or transfected with a DNA construct according to claim 3.
- 5. A cell stably transformed or transfected with a DNA construct according to claim 4.

\* \* \* \* \*